

Caucasus — the mountain of languages

Manana Tandashvili
University of Frankfurt

The widespread picture of linguistic diversity in the Caucasus as ‘the mountain of languages’ will be immediately confirmed if a closer look is taken at the region: multiethnic, multilingual, multireligious is the adequate description of this melting pot. What is responsible for the present-day ethnic, linguistic and sociocultural diversity is the historical coexistence of different ethnic groups in a geographically delimited region on the one hand, and the geopolitical situation at the border between the Orient and the Occident on the other. At the same time, this diversity leads to mutual influence of different kinds, ranging from linguistic and religious to ethnic assimilation. In this article, we will outline the results of relevant international projects in the field of ‘language documentation’ that we conducted over the past 15 years and what we have learned from these projects.

1. The study of the Caucasian Languages Research on Caucasian languages has been going on for quite some time. Even though Caucasian linguistics emerged a long time later than Indo-European, Semitic, or Uralian studies, it has a longer tradition than most other areas. Of the 70 odd languages that are spoken in the Caucasus, only Georgian can look back on centuries of uninterrupted written tradition, thus pertaining to the best documented languages worldwide, in contrast to other Caucasian languages whose written materials date only from the end of the Middle Ages (e.g., for Avar from the 14th, for Darginian from the 16th or for Tabasaran from the 17th century). All these materials mostly consist of word lists (compiled by scholars such as Güldenstädt, Klaproth, Peacock and others), which were often incorrectly transcribed because of a lack of knowledge of the Caucasian languages and are mostly unreliable as research materials.

Intensive studies of Caucasian languages began in the second half of the 19th century, commissioned by the Russian Academy of Sciences. Investigations by Peter von Uslar, Anton Schiefner, Adolf Dirr and others brought about numerous grammars of Caucasian languages such as Abkhaz, Chechen, Avar, Udi, and others. In addition to a description

of the language structure, these grammars often include concise dictionaries and texts of the respective languages.

In the Soviet era, the study of the Caucasian languages took a systematic and planned character: many languages, especially undocumented ones, were systematically recorded, fieldwork was conducted regularly, dictionaries were created and texts were published.

The Soviet stage of Caucasian language research was by all means a step forward in the documentation of Caucasian languages, but the scope of these efforts remained limited due to several factors such as the restricted personal competence of both speakers and scholars. The methods of recording the relevant materials were very differentiated and often inaccurate. Most of the time the people that were chosen as native speakers had a good competence of the language indeed, but the dynamics of the language situation was not considered, linguistic varieties being documented without any sociolinguistic background information (as to, e.g., language proficiency by age or social status of speakers).

In the 21st century, the development of technologies in language processing have completely changed the basis for documenting languages. In addition, theoretical approaches have been conceived that have facilitated new approaches, with Nikolaus Himmelmann's article of 1998 marking a decisive step forward, ensuring that Documentary Linguistics was institutionalized as a new subject of linguistic research, with a high impact on the Caucasian languages, too.

2. Endangered Caucasian languages in Georgia When the Volkswagen Foundation launched the call for proposals of the program DOBES ("Dokumentation bedrohter Sprachen" / "Documentation of endangered languages"), we developed a project at the Institute for Comparative Linguistics¹ of the University of Frankfurt, which aimed at identifying the endangered languages in one of the South Caucasian countries. We chose three languages spoken in Georgia to document and determine their degree of endangerment:

1. Udi, an East Caucasian language belonging to the Lezgic group, which is spoken in Georgia but also in Azerbaijan and elsewhere;
2. Batsbi or Tsova-Tush, an East Caucasian language belonging to the Nakh group, which is spoken only in Georgia, more precisely in just one village;
3. Svan, one of the South Caucasian languages, which is spoken mainly in Svanetia (in the Georgian high mountain area) but, because of recent migration, also in small groups (linguistic islands) in the lowlands of Eastern Georgia.

The DOBES project ECLinG (Endangered Caucasian Languages in Georgia), which dealt with the three aforementioned languages, was conducted in 2002-2006 in cooperation with Georgian partner institutions (the A. Chikobava Institute of Linguistics and the I. Javakhishvili State University, Tbilisi).

At the beginning of the project, when we started to record speakers of our object languages, we realised how strongly the minority language communities are characterised by multilingualism and what impact this has on the endangerment of their language. The research group fully shared N. Himmelmann's view that in the case of highly endangered languages, it is particularly important to document them, as further data collection in the

¹Since 2010 "Institut für Empirische Sprachwissenschaft" [Institute for Empirical Linguistics].

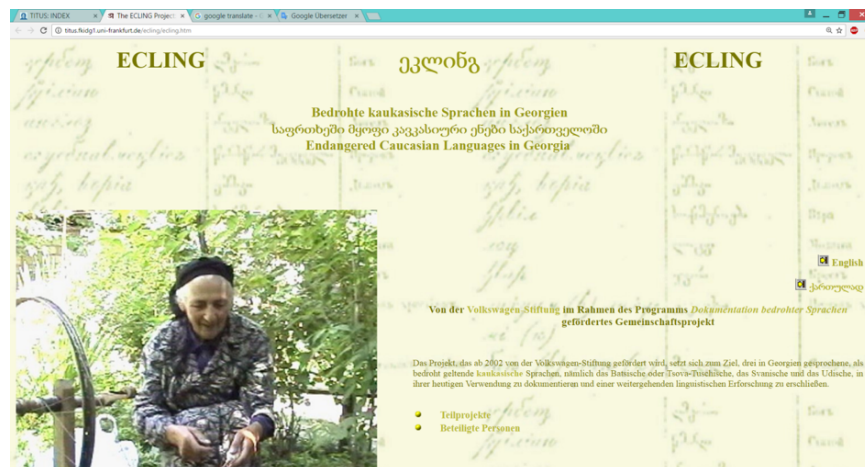


Figure 1: Screenshot of ECLing project website

future may be difficult if not impossible (Himmelman, Nikolaus P. 1998). For this reason, after the completion of the ECLInG project, we continued to document the small language communities in Georgia and continued our work in the new project “Sociolinguistic Situation in Contemporary Georgia” (SSGG; 2006-2009). The data collected in both projects are available in the TLA archive at the MPI for Psycholinguistics in Nijmegen.² In addition, the data have also been integrated into the databases of the TITUS project³ and the Georgian National Corpus (GNC). The results of our linguistic and sociolinguistic research have been discussed in various books and article (Chantladze et al. 2007–2010; Gippert 2008; Gippert 2012; Aristar-Dry et al. 2012; Tandashvili 2011; Tandashvili 2015).

For both scientific projects, our task was divided into two parts: language documentation and language description. In the following, I will discuss these two aspects using the example of two minorities in Georgia, the Udi and the Laz people.

3. Language description and (socio-) linguistic observations Within the ECLInG project, we were able to describe the language system of the three target languages and compile dictionaries of these languages. The text materials were first edited in Shoebox, later in Toolbox, with transcriptions, linguistic annotations, and translations into Georgian and English. In total, we have processed more than 70 hours of recordings in the ECLInG project.

On the basis of the audio-video recordings we hope to have prepared a solid base for research into all areas of the phonology, morphology, syntax etc. of these languages, but also their cultural properties. For example, we covered topics such as a) aspects of Christian and pre-Christian religion, b) material culture determined by a peculiar environment (the high-mountainous setting), with a focus on sheep breeding (in Batsbi), spinning, weaving and knitting (also in Batsbi), cattle breeding (in Svan), agriculture (in Svan and Udi), and wine growing (in Udi), but also c) folklore in form of singing and dancing (for Svan and Batsbi).

²The Language Archive. (<https://archive.mpi.nl>)

³Thesaurus Indogermanischer Text- und Sprachmaterialien (<http://titus.uni-frankfurt.de>)

While analyzing the recorded texts linguistically, we came across an interesting phenomenon that we call fluctuating or situational competence. This phenomenon is characterized by a different speech behaviour of a given speaker and shows different levels of competence depending on the situation: in a monologue the speaker cannot remember a certain word, changes the code and continues the talk in another (usually the dominant) language or uses foreign words from this language, while the corresponding lexical items are readily present in a natural (dialogic) speech situation—the word is ‘activated’ by this context and thus used without hesitation.

Often, for example, Udi speakers are not familiar with specific agricultural terms. In this case they use Georgian lexemes, integrating them in accordance with the verbal morphology of Udi, but at the same time they are aware of this process and sometimes explicitly signal it as being a change of code:

- (1) *Oša me čapnux gapurčna-yan-besa...*
 ‘Then we thin this vineyard out...’ (in Udi)
es ukve kartuli siṭqvaa.
 ‘this is already a Georgian word’ (in Georgian)

It is further important that code-switching is a natural thing in the multilingual environment of the Udis, often provoked by the incompetence of the speaker:

- (2) *Tene hebsa, davateyanduxsa, ič benginänne, täksä heba... gasxvla rogor aris beš muzin, tezaaba.*
 ‘Nothing is necessary, we do not disinfect, it grows by itself, alone ... What *gasxvla* (trim the vines) is in our language, I don’t know.’

However, as the next example shows, this ‘incompetence’ can be overcome if the speaker uses the word in a natural context (here during the vintage in October):

- (3) *Siftä kaçyanexa čapnu, toxyanbesa, davyanduxsa.*
 ‘First we cut the vine, we weed it, we disinfect it.’

Linguistic behavior in extra-contextual cases is often accompanied by code switching. Based on our materials, we have divided the manifestations of code switching into two types:

1. ‘signalled’ code switching, i.e. code switching that is caused by triggering words and is often associated with incompetence of the speakers;
2. ‘non-signalled’ code switching, which is exerted between two languages without a signal and without damaging the ecology of the language. The switching of the code can e.g. be carried out in the middle of a subordinate clause as in example (4); here the boundary between the two languages and the syntactic boundaries in the sentence do not coincide.

- (4) ... { [*gasxvla rogor aris*]_{Georgian} [*beš muzin, }_{SUB} tezaaba*]_{Udi}
 [what ‘trimming’ is]_{Georgian} [in our language, I don’t know]_{Udi}

With Udi speakers, code switching can be observed not only within sentences but also within numerical expressions. In our recordings we have documented cases where ones and tens are composed using components from different languages, as in *qər biṗ* ‘forty-four’, composed of Azeri *qər* and Udi *biṗ*.

A particularly interesting issue of the project was to compare the language situation of the two ‘insular’ varieties of Svan in the Kodori valley and in the lowlands. While Kodori Svan does not show any specific changes in the grammatical system or in the lexicon, the lowland variety seems to be far more endangered, although in both cases the children attend Georgian schools and their education as well as the available media use Georgian only. This indicates that the geographic seclusion of a linguistic island with little contact with the main language area is not necessarily the primary cause of increasing endangerment. Rather, the variety in question can evolve with its own language structure and, in extreme cases, end as an independent dialect. However, an island whose sociolinguistic context is determined by a dominant surrounding language can show within two generations all the characteristics of endangerment, which manifest themselves in the abolition of the continuous transmission of language and cultural heritage.

Among the Svan people now living in the eastern Georgian lowlands, we have seen yet another phenomenon, namely a special case of multilingualism, which was also observed later on in the Laz community in the framework of the SSGG project. In the conversation between the older and younger generation often two languages are used: the grandfather speaks with the grandchild in Svan or Laz; the child understands what has been said but cannot answer in the same language and so uses Georgian. Thus the entire dialogue takes place in two languages, and each participant can be assigned one of them.

Comparing the linguistic situation of the Batsbi people in Georgia with that of the Udi and Svan people, one finds another case of assimilation, which will be discussed in the following section.

4. Collecting data for multidisciplinary research As the examples presented above show, the data collected in our project can serve not only as the basis for (socio)linguistic studies but also provide much material for other disciplines such as ethnology, ethnopsychology, ethnography, oral literature, religious history, conflict research, etc.

In order to describe the language situation more accurately and to document the Udi language in its fields of use, we tried to observe and continuously document Udi speakers for more than 10 years, visiting the locations at least three times a year—in spring (March–April), in summer (July–August) and in autumn (in October).

While the first recordings were still relatively spontaneous and unplanned, aiming to become acquainted with important topics such as personal biographies, household, everyday activities or fields of employment and to collect the relevant vocabulary, in later years we took recordings with more specific goals. So, we filmed Udi people e.g. during the vintage or went to a wedding in order to document the natural process of this ritual. In 2011, we attended the inauguration of the first church in Zinobiani asking the Udi people about the importance of religion with regard to their (cultural) identity.

All these recordings show a remarkable asymmetry between the various components of identity (linguistic, ethnic, and religious), which I studied in my essay “What causes the endangerment of languages? The Case of the Udi Language in Georgia” (Tandashvili, Manana, 2018). The exploration of the language situation of the Udi and the Batsbi

people has clearly shown that the linguistic identity does not necessarily have to be linked to an ethnic identity, just as the ethnic identity does not necessarily derive from a linguistic identity. Thus, the Udi people are almost completely assimilated linguistically, but ethnically they have maintained their identity. The Batsbi people, conversely, are assimilated ethnically by identifying themselves as Georgians but have kept their identity linguistically. This results in a particular identity constellation: the terms 'Batsbi people' and 'Svan people' cannot be used as ethnic terms but refer to the linguistic identity of the community concerned, while 'Udi people' in Georgia characterizes an ethnic group.

Assuming that the ethnic identity is based on 'historical memories', the 'ethnic stability' of the Udi people can well be understood. However, it is also known that this type of collective memory can be forgotten, revitalized, or reshaped. The 'historical memory' of the Batsbi people is largely erased, while the Udi people have revitalized theirs by referring to the medieval state of 'Albania' in the Caucasus. The discovery of the Caucasian Albanian manuscripts in St. Catherine's Monastery on Sinai, on the one hand, and the idea of being descendants of one of the oldest people in the southern Caucasus, on the other, today plays a major role in the cultural identity of the Udi people. Whether the official recognition of their ethnicity has helped here remains unanswered and requires further investigation. About the Batsbi people, however, it can be said by now that their 'historical memory' is deleted.

References

- Aristar-Dry, Helen, Sebastian Drude, Menzo Windhouwer, Jost Gippert & Irana Nevskaya. 2012. Rendering Endangered Lexicons Interoperable through Standards Harmonization: the RELISH Project. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*, 766–770. Istanbul: European Language Resources Association.
- Chantladze, Iza, Ketevan Margiani-Dadvani, Ketevan Margiani-Subari, Medea Sagliani & Rusudan Ioseliani. 2007–2010. *The Kodori Chronicles. (With Researches). Vol. I*. Tbilisi.
- Gippert, Jost. 2008. Endangered Caucasian Languages in Georgia. In David K. Harrison, David S. Rood & Arienne Dwyer (eds.), *Lessons from Documented Endangered Languages*, 159–194. (Typological Studies in Languages 78). Amsterdam: John Benjamins.
- Gippert, Jost. 2012. Language-specific encoding in endangered language corpora. In Frank Seifart, Geoffrey Haig, Nikolaus P. Himmelmann, Dagmar Jung, Anna Margetts & Paul Trilsbeek (eds.), *Potentials of Language Documentation: Methods, Analyses, and Utilization*, 17–24. (Language Documentation & Conservation Special Publication, 3). Honolulu: University of Hawai'i Press.
- Himmelmann, Nikolaus P. 1998. Documentary and Descriptive Linguistics. *Linguistics* 36. 161–195.
- Tandashvili, Manana. 2009: Georgien – multiethnisch, multilingual, multireligiös. *PRO GEORGIA. Journal of Kartvelological Studies* 18. 51–65.
- Tandashvili, Manana. 2011. Change of Codes under mulilingual conditions and idiolectal Interferences. In Iza Chantladze, Eter Soselia, Nugsar Mgeladze (eds.), *Linguocultural Researches*, 161–171. Batumi: Shota Rustaveli State University Press.
- Tandashvili, Manana. 2015: Sprachliche Minderheiten in Georgien (Zur Methodologie). In Natia Reineck & Ute Rieger (eds.), *Kaukasiologie heute. Festschrift für Heinz Fähnrich zum 70. Geburtstag*, 395–410. Jena.
- Tandashvili, Manana. 2018. What causes the endangerment of languages? The Case of the Udi Language in Georgia. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 15. 289–312.

Digital Resources

The Language Archive

(<https://archive.mpi.nl>)

Thesaurus Indogermanischer Text- und Sprachmaterialien

(<http://titus.uni-frankfurt.de>)

ECLinG - Endangered Caucasian Languages in Georgia

(<http://titus.fkidg1.uni-frankfurt.de/ecling/ecling.htm>)

SSGG - Die soziolinguistische Situation im gegenwärtigen Georgien

(<http://titus.uni-frankfurt.de/ssgg/ssgg.htm>)

BaLDAR Batumi Linguocultural Digital Archive

(<http://digiarchive.bsu.edu.ge/>)

Manana Tandashvili

tandaschwili@em.uni-frankfurt.de